OS90/1108

#2   OIPE

RAW SEQUENCE LISTING   DATE: 10/18/2001
PATENT APPLICATION: US/09/966,608   TIME: 17:10:16

Input Set : A:\PM4966.txt
Output Set: N:\CRF3\10182001\I966608.raw   **ENTERED**

```
 4 <110> APPLICANT: Braun, Jonathan
 5         Sutton, Christopher L.
 7 <120> TITLE OF INVENTION: IBD-Associated Microbial Nucleic Acid
 8         Molecules
10 <130> FILE REFERENCE: P-PM 4966
12 <140> CURRENT APPLICATION NUMBER: US/09/966,608      C-->
12 <141> CURRENT FILING DATE: 2001-09-27               C-->
12 <150> PRIOR APPLICATION NUMBER: US 09/303,120
13 <151> PRIOR FILING DATE: 1999-04-30
15 <150> PRIOR APPLICATION NUMBER: US 09/820,576
16 <151> PRIOR FILING DATE: 2001-03-28
18 <160> NUMBER OF SEQ ID NOS: 10
20 <170> SOFTWARE: FastSEQ for Windows Version 4.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 302
24 <212> TYPE: DNA
25 <213> ORGANISM: Unknown
27 <220> FEATURE:
28 <223> OTHER INFORMATION: Microbial Organism from the human gut
30 <221> NAME/KEY: CDS
31 <222> LOCATION: (2)...(301)
33 <400> SEQUENCE: 1
34 a gat ctg gcc agc gcc gtg ggc atc cag tcc ggc agc atc ttt cat cac 49
35   Asp Leu Ala Ser Ala Val Gly Ile Gln Ser Gly Ser Ile Phe His His
36    1               5                  10                  15
38 ttc aag agc aag gat gag ata ttg cgt gcc gtg atg gag gaa acc atc   97
39 Phe Lys Ser Lys Asp Glu Ile Leu Arg Ala Val Met Glu Glu Thr Ile
40            20                  25                  30
42 cat tac aac acc gcg atg atg cgc gct tca ctg gag gag gcg agc acg   145
43 His Tyr Asn Thr Ala Met Met Arg Ala Ser Leu Glu Glu Ala Ser Thr
44            35                  40                  45
46 gtg cgc gaa cgc gtg ctg gcg ctg atc cgc tgc gag ttg cag tcg atc   193
47 Val Arg Glu Arg Val Leu Ala Leu Ile Arg Cys Glu Leu Gln Ser Ile
48       50                  55                  60
50 atg ggc ggc agt ggc gag gcc atg gcg gtg ctg gtc tac gaa tgg cgc   241
51 Met Gly Gly Ser Gly Glu Ala Met Ala Val Leu Val Tyr Glu Trp Arg
52   65                  70                  75                  80
54 tcg ctg tcg gcc gaa ggc cag gcg cac gtg ctg gcc ctg cgt gac gtg   289
55 Ser Leu Ser Ala Glu Gly Gln Ala His Val Leu Ala Leu Arg Asp Val
56               85                  90                  95
58 tat gag cag atc t                                               302
59 Tyr Glu Gln Ile
60            100
63 <210> SEQ ID NO: 2
64 <211> LENGTH: 100
65 <212> TYPE: PRT
66 <213> ORGANISM: Unknown
```

```
68 <220> FEATURE:
69 <223> OTHER INFORMATION: Microbial organism from the human gut
71 <400> SEQUENCE: 2
72 Asp Leu Ala Ser Ala Val Gly Ile Gln Ser Gly Ser Ile Phe His His
73 1               5                   10                  15
74 Phe Lys Ser Lys Asp Glu Ile Leu Arg Ala Val Met Glu Glu Thr Ile
75              20                  25                  30
76 His Tyr Asn Thr Ala Met Met Arg Ala Ser Leu Glu Glu Ala Ser Thr
77          35                  40                  45
78 Val Arg Glu Arg Val Leu Ala Leu Ile Arg Cys Glu Leu Gln Ser Ile
79      50                  55                  60
80 Met Gly Gly Ser Gly Glu Ala Met Ala Val Leu Val Tyr Glu Trp Arg
81 65                  70                  75                  80
82 Ser Leu Ser Ala Glu Gly Gln Ala His Val Leu Ala Leu Arg Asp Val
83                  85                  90                  95
84 Tyr Glu Gln Ile
85                 100
88 <210> SEQ ID NO: 3
89 <211> LENGTH: 392
90 <212> TYPE: DNA
91 <213> ORGANISM: Unknown
93 <220> FEATURE:
94 <223> OTHER INFORMATION: Microbial Organism from the human gut
96 <221> NAME/KEY: CDS
97 <222> LOCATION: (2)...(346)
99 <221> NAME/KEY: misc_feature
100 <222> LOCATION: (1)...(392)
101 <223> OTHER INFORMATION: n = A,T,C or G
103 <400> SEQUENCE: 3
104 a gat ctt gag cgt cat gag tgc ctg ggg tac gcc ttt tca tcg cgt ccg 49
105   Asp Leu Glu Arg His Glu Cys Leu Gly Tyr Ala Phe Ser Ser Arg Pro
106   1               5                   10                  15
108 gcg gat cga gag tgg gtg ttt ttt cag ggc acg gtt tcc tac aag gta 97
109 Ala Asp Arg Glu Trp Val Phe Phe Gln Gly Thr Val Ser Tyr Lys Val
110              20                  25                  30
112 cga gtg gcc agc cgt ttg ctc atc aat gaa agc cgg gca ttg atg tcg 145
113 Arg Val Ala Ser Arg Leu Leu Ile Asn Glu Ser Arg Ala Leu Met Ser
114              35                  40                  45
116 gcg gca ttg gat ggt ttt ggc ata gtg ctc ggc ccg caa gac ttc ctg 193
117 Ala Ala Leu Asp Gly Phe Gly Ile Val Leu Gly Pro Gln Asp Phe Leu
118      50                  55                  60
120 cga acg gcg ttg gcg agt ggc gag ttg gtg cgg gtg ttg ccg gag ttt 241
121 Arg Thr Ala Leu Ala Ser Gly Glu Leu Val Arg Val Leu Pro Glu Phe
122 65                  70                  75                  80
124 gag gct ccg agt cgg tcg atg cat ttg gtc tac acc gca aac cgc cag 289
125 Glu Ala Pro Ser Arg Ser Met His Leu Val Tyr Thr Ala Asn Arg Gln
126                  85                  90                  95
128 cgt acc gcc aag ttg cgc tgc ttt gtc gag act gtg ctg gga cgt ttt 337
129 Arg Thr Ala Lys Leu Arg Cys Phe Val Glu Thr Val Leu Gly Arg Phe
```

RAW SEQUENCE LISTING   DATE: 10/18/2001
PATENT APPLICATION: US/09/966,608   TIME: 17:10:16

Input Set : A:\PM4966.txt
Output Set: N:\CRF3\10182001\I966608.raw

```
      130             100             105             110
W-->  132 ggt ccg gta tgaaggagca ccaccgtggc ggtcgccggg angcacctaa          386
      133 Gly Pro Val
      134           115
      136 agatct                                                           392
      138 <210> SEQ ID NO: 4
      139 <211> LENGTH: 115
      140 <212> TYPE: PRT
      141 <213> ORGANISM: Unknown
      143 <220> FEATURE:
      144 <223> OTHER INFORMATION: Microbial organism from the human gut
      146 <400> SEQUENCE: 4
      147 Asp Leu Glu Arg His Glu Cys Leu Gly Tyr Ala Phe Ser Ser Arg Pro
      148 1               5                   10                  15
      149 Ala Asp Arg Glu Trp Val Phe Phe Gln Gly Thr Val Ser Tyr Lys Val
      150           20                  25                  30
      151 Arg Val Ala Ser Arg Leu Leu Ile Asn Glu Ser Arg Ala Leu Met Ser
      152           35                  40                  45
      153 Ala Ala Leu Asp Gly Phe Gly Ile Val Leu Gly Pro Gln Asp Phe Leu
      154      50                  55                  60
      155 Arg Thr Ala Leu Ala Ser Gly Glu Leu Val Arg Val Leu Pro Glu Phe
      156 65                  70                  75                  80
      157 Glu Ala Pro Ser Arg Ser Met His Leu Val Tyr Thr Ala Asn Arg Gln
      158                 85                  90                  95
      159 Arg Thr Ala Lys Leu Arg Cys Phe Val Glu Thr Val Leu Gly Arg Phe
      160           100                 105                 110
      161 Gly Pro Val
      162           115
      165 <210> SEQ ID NO: 5
      166 <211> LENGTH: 114
      167 <212> TYPE: PRT
      168 <213> ORGANISM: Unknown
      170 <220> FEATURE:
      171 <223> OTHER INFORMATION: Microbial Organism from the human gut
      173 <221> NAME/KEY: VARIANT
      174 <222> LOCATION: (1)...(114)
      175 <223> OTHER INFORMATION: Xaa = Any Amino Acid
      177 <400> SEQUENCE: 5
      178 Arg Thr Arg Arg Ile Ser Leu Pro His Lys Lys Leu Ala Arg Asn Gly
      179 1               5                   10                  15
      180 Val Leu Tyr Ser His Gly Ala Thr Gln Glu Asp Ile Phe Ala Pro Cys
      181           20                  25                  30
      182 Gln His Arg Arg Cys Gln Ile Thr Lys Ala Tyr His Glu Ala Arg Leu
      183           35                  40                  45
      184 Val Glu Gln Ser Arg Arg Gln Arg Thr Ala Leu Gln His Pro His Gln
      185      50                  55                  60
      186 Arg Leu Lys Leu Ser Arg Thr Pro Arg His Met Gln Asp Val Gly Cys
      187 65                  70                  75                  80
      188 Val Ala Leu Thr Gly Gly Leu Gln Ala Ala Lys Asp Leu Ser His Gln
```

RAW SEQUENCE LISTING   DATE: 10/18/2001
PATENT APPLICATION: US/09/966,608   TIME: 17:10:16

Input Set : A:\PM4966.txt
Output Set: N:\CRF3\10182001\I966608.raw

```
189                     85                  90                  95
190 Ser Thr Lys Thr Arg Tyr Ser Pro Ala Gly Gly His Arg Asp Gly Pro
191             100                 105                 110
```
W--> 192 Xaa Val
```
196 <210> SEQ ID NO: 6
197 <211> LENGTH: 190
198 <212> TYPE: PRT
199 <213> ORGANISM: Clostridium pasteurianum
201 <400> SEQUENCE: 6
202 Met Asn Lys Thr Lys Asp Asn Ile Phe Tyr Ser Ala Ile Lys Val Phe
203 1               5                   10                  15
204 Ser Asn Asn Gly Tyr Asn Gly Ala Thr Met Asp Glu Ile Ala Ser Asn
205             20                  25                  30
206 Ala Gly Val Ala Lys Gly Thr Leu Tyr Tyr His Phe Lys Ser Lys Glu
207             35                  40                  45
208 Glu Ile Phe Lys Tyr Ile Ile Glu Glu Gly Val Asn Leu Met Lys Asn
209     50                  55                  60
210 Glu Ile Asp Glu Ala Thr Asp Lys Glu Lys Thr Ala Leu Glu Lys Leu
211 65                  70                  75                  80
212 Lys Ala Val Cys Arg Val Gln Leu Asn Leu Ile Tyr Lys Asn Arg Asp
213             85                  90                  95
214 Phe Phe Lys Val Ile Ala Ser Gln Leu Trp Gly Lys Glu Leu Arg Gln
215             100                 105                 110
216 Leu Glu Leu Arg Asp Ile Met Arg Asn Tyr Val Val His Ile Glu Glu
217             115                 120                 125
218 Phe Val Lys Asp Ala Met Glu Ala Gly Ser Ile Lys Lys Gly Asn Ser
219     130                 135                 140
220 Leu Phe Val Ala Tyr Ala Phe Leu Gly Thr Leu Cys Ser Val Ser Leu
221 145                 150                 155                 160
222 Tyr Glu Val Ile Asn Ala Glu Asn Asp Asn Ile Asn Asn Thr Ile Glu
223             165                 170                 175
224 Asn Leu Met Asn Tyr Ile Leu Asn Gly Ile Gly Leu Gln Asn
225             180                 185                 190
228 <210> SEQ ID NO: 7
229 <211> LENGTH: 200
230 <212> TYPE: PRT
231 <213> ORGANISM: Mycobacterium tuberculosis
233 <400> SEQUENCE: 7
234 Met Asp Arg Val Ala Gly Gln Val Asn Ser Arg Arg Gly Glu Leu Leu
235 1               5                   10                  15
236 Glu Leu Ala Ala Ala Met Phe Ala Glu Arg Gly Leu Arg Ala Thr Thr
237             20                  25                  30
238 Val Arg Asp Ile Ala Asp Gly Ala Gly Ile Leu Ser Gly Ser Leu Tyr
239             35                  40                  45
240 His His Phe Ala Ser Lys Glu Glu Met Val Asp Glu Leu Leu Arg Gly
241     50                  55                  60
242 Phe Leu Asp Trp Leu Phe Ala Arg Tyr Arg Asp Ile Val Asp Ser Thr
243 65                  70                  75                  80
244 Ala Asn Pro Leu Glu Arg Leu Gln Gly Leu Phe Met Ala Ser Phe Glu
```

RAW SEQUENCE LISTING   DATE: 10/18/2001
PATENT APPLICATION: US/09/966,608   TIME: 17:10:16

Input Set : A:\PM4966.txt
Output Set: N:\CRF3\10182001\I966608.raw

```
245                   85                  90                  95
246 Ala Ile Glu His His His Ala Gln Val Val Ile Tyr Gln Asp Glu Ala
247             100                 105                 110
248 Gln Arg Leu Ala Ser Gln Pro Arg Phe Ser Tyr Ile Glu Asp Arg Asn
249         115                 120                 125
250 Lys Gln Gln Arg Lys Met Trp Val Asp Val Leu Asn Gln Gly Ile Glu
251     130                 135                 140
252 Glu Gly Tyr Phe Arg Pro Asp Leu Asp Val Asp Leu Val Tyr Arg Phe
253 145                 150                 155                 160
254 Ile Arg Asp Thr Thr Trp Val Ser Val Arg Trp Tyr Arg Pro Gly Gly
255             165                 170                 175
256 Pro Leu Thr Ala Gln Gln Val Gly Gln Gln Tyr Leu Ala Ile Val Leu
257             180                 185                 190
258 Gly Gly Ile Thr Lys Glu Gly Val
259         195                 200
262 <210> SEQ ID NO: 8
263 <211> LENGTH: 192
264 <212> TYPE: PRT
265 <213> ORGANISM: Auifex aeolicus
267 <400> SEQUENCE: 8
268 Met Tyr Ile Leu Leu Phe Met Gly Glu Lys Arg Ser Asp Thr Lys Glu
269 1                   5                   10                  15
270 Lys Ile Leu Ser Ser Ala Leu Lys Leu Phe Ser Lys Lys Gly Phe Lys
271             20                  25                  30
272 Glu Thr Thr Ile Lys Asp Ile Ala Lys Glu Val Gly Ile Thr Glu Gly
273         35                  40                  45
274 Ala Ile Tyr Arg His Phe Thr Ser Lys Glu Glu Ile Ile Lys Ser Leu
275     50                  55                  60
276 Leu Glu Ser Ile Thr Lys Glu Leu Arg His Lys Leu Glu Val Ala Leu
277 65                  70                  75                  80
278 Gln Arg Gly Glu Thr Asp Glu Glu Ile Leu Glu Ser Ile Val Asp Thr
279                 85                  90                  95
280 Leu Ile Asp Tyr Ala Phe Ser Asn Pro Glu Ser Phe Arg Phe Leu Asn
281             100                 105                 110
282 Leu Tyr His Leu Leu Lys Glu Tyr Gly Glu Val Lys Asn Leu Pro Gly
283         115                 120                 125
284 Glu Leu Ile Leu Lys Phe Leu Asn Gly Leu Tyr Leu Lys Arg Lys Leu
285     130                 135                 140
286 Lys Thr Tyr Pro Glu Ile Ala Leu Ala Val Val Thr Gly Ser Val Glu
287 145                 150                 155                 160
288 Arg Val Phe Ile Phe Lys Glu Arg Asn Phe Leu Asp Tyr Asp Glu Glu
289             165                 170                 175
290 Thr Ile Lys Lys Glu Leu Lys Lys Val Leu Lys Ser Ala Ile Leu Ala
291             180                 185                 190
294 <210> SEQ ID NO: 9
295 <211> LENGTH: 18
296 <212> TYPE: DNA
297 <213> ORGANISM: Unknown
299 <220> FEATURE:
```

VERIFICATION SUMMARY   DATE: 10/18/2001
PATENT APPLICATION: US/09/966,608   TIME: 17:10:17

    Input Set : A:\PM4966.txt
    Output Set: N:\CRF3\10182001\I966608.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:132 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:192 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5